This relationship has important implications and deserves further study.

It should be noted that while regression analysis is a useful method for ranking the importance of a series of variables as they impact a "dependent" variable, there are certain potential limitations that must be kept in mind. While regression can show that an association exists, it is not safe to conclude causality just on this basis. Though the results may be suggestive, regression analysis **alone** does not demonstrate that a change in one independent variable directly leads to change in the dependent one. Regression may demonstrate coexistence, but causality must be imputed on theoretrical grounds or demonstrated by experimental methods.

Linear multiple regression, the type employed here, also involves certain assumptions; and if serious departures from these assumptions occur, the results may be biased. First, each of the dependent variables was plotted and found to approximate a normal distribution. We tested the assumption of a linear relationship between each independent variable and total discharges per 1,000 population by plotting the pairs of variables and found no serious deviations from linearity. Other assumptions are that the residual or error term (actual value minus predicted value) is uncorrelated with other variables in the equation and that the error variance is constant across different values of the other variables. Plots of the residuals against the predicted values, the sequence of cases, and each independent variable revealed that these assumptions were not violated.

In an earlier study in this series (5), a suggested approach was to predict hospital use rates based on demographic and need factors and then to determine what other variables account for deviations from these expected "need" values. As the present study developed, the difficulty of isolating one set of variables and saying a priori that **these** indicate need became more apparent. This is particularly true given the sometimes high correlations between the residence-based variables and the medical resource variables. Thus no attempt was made here to isolate a set of need indicators, but rather all major variables were included in the regression model to compare their importance, and deviations from the resulting predicted values were examined. The approach of quantitatively predicting a "needed" level of hospital use should, however, be pursued in future research.

While our analysis accounted for around 60 percent of the variation in hospital utilization, 40 percent remains unexplained by the variables that we could measure in a county-level analysis. One important variable not in this study is health insurance coverage. It has been well documented that lack of health insurance coverage is associated with low hospital utilization rates (6,15,22); therefore, a measure of the percent of county residents without hospital insurance could add to our ability to predict hospital utilization. Another variable that impacts hospital utilization is accessibility of hospital care (14). Dimensions of accessibility would include availability, travel distance, ability to pay, and acceptability of the services offered. Variations from county to county in the health status of the population should also account for some of this unexplained variation, though this dimension would be very expensive to quantify.

In conclusion, it is hoped that the results of this study will be useful to health planners and those involved in the delivery of health care in a county. Examining a county's values on variables found important in the prediction of hospital use may help explain unusually high or low levels of use. Furthermore, for counties with a much higher or lower level of use than would be expected given its values on the predictor variables (i.e., for counties with large residuals), further investigation may uncover reasons for these unusual patterns of hospital utilization that we were not able to measure. It should, however, be noted that even though a county's hospital utilization rate deviates from an average or expected rate, it does not necessarily mean that something is wrong. A low utilization rate could be due to less than optimal utilization, but it could also be due to very little unnecessary utilization or low morbidity. On the other hand, a high utilization rate could result from a high level of morbidity in that county. Questions about unusually high or low utilization can best be answered by those familiar with the situation in a particular county.